Validation of a Standardised Duplex Ultrasound Classification System for the Reporting and Grading of Peripheral Arterial Disease

Samuel Huthart a,b,c,*, Crispian Oates a,b, John Allen a,b,d, Kathia Fiaschi a,b,c, Andrew J. Sims b, Gerard Stansby a,b

a Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK
b Northern Medical Physics and Clinical Engineering Department, Freeman Hospital, Newcastle upon Tyne, UK
c Northern Vascular Centre, Freeman Hospital, Newcastle upon Tyne, UK
d Research Centre for Intelligent Healthcare, Coventry University, Coventry, UK

WHAT THIS PAPER ADDS
This study used a simplified and standardised report of the duplex ultrasound (DUS) examination in peripheral arterial disease (PAD), independent of any need for images or exact measurements to be recorded. Simplification and standardisation of the process should offer efficiencies in both reporting and documentation of PAD, as well as its severity in the non-specialist/community setting following initial screening methods, or for research in PAD diagnostics, clinical trials, or when intervention is being considered.

Objective: Duplex ultrasound (DUS), a non-invasive means of arterial mapping, allows for the reliable diagnosis of peripheral arterial disease (PAD). One of the authors (C.P.O.), developed a standardised DUS based scoring system, devised for rapid detection and reporting of PAD. The purpose of this study was to validate this system, and to determine the diagnostic performance both overall and per disease severity.

Methods: In total, 250 participants were recruited, based on diagnosis of (n = 125) or absence of PAD (n = 125) from general practice registers. Right and left legs per subject were handled as independent readings, determining actual PAD status via ankle brachial pressure index (ABPI) < 0.9, and then further grading disease severity using suggested ABPI ranges. Data were excluded if no corresponding ABPI value was obtained per DUS determination or if the ABPI reading was > 1.4, owing to the risk of false negatives due to incompressible vessels. Diagnostic sensitivity and specificity were obtained overall, and per severity classification. Furthermore, inter-rater agreement between ABPI and DUS determined PAD severity was determined by linear weighted Cohen’s kappa.

Results: The sensitivity and specificity in the detection of disease overall was 81.0% (95% confidence interval [CI] 73.4 — 87.2) and 86.3% (95% CI 82.3 — 89.8), respectively. From mild to severe PAD, sensitivity increased from 71.1% (95% CI 55.7 — 83.6) to 89.3% (95% CI 71.8 — 97.7). Furthermore, a Cohen’s kappa value of 0.63 (95% CI 0.57 — 0.69) was obtained, indicating moderate agreement between the two diagnostic methods.

Conclusion: The findings of this study validate the diagnostic performance of the standardised DUS scoring system, as well as its capacity to grade severity of disease, offering a potential tool for the identification of PAD in community/research settings following initial screening methods. Confirmatory work could include a comparison of DUS determined disease with gold standard methods of non-invasive angiography, and novel tools such as toe flex near infrared spectroscopy and multisite photoplethysmography.

Keywords: Ankle brachial pressure index, Doppler waveform, Duplex ultrasound, Peripheral arterial disease, Vascular diagnostics

INTRODUCTION
Peripheral arterial disease (PAD) is a major cardiovascular disease believed to affect approximately > 200 million people worldwide.1 It is currently the third most common clinical manifestation of atherosclerosis,2 and despite the significant associated cardiovascular morbidity and mortality,1,3 it remains largely underdiagnosed and under-managed.1,4 Furthermore, the incidence increases with age, and given the ageing population (particularly in middle to high income countries), PAD is likely to become of even greater concern to public health in the future.4

Duplex ultrasound (DUS) is a widely used non-invasive means of peripheral arterial mapping that has demonstrated median diagnostic sensitivities of 88% (95% confidence interval [CI] 80 — 98) and specificities of 96% (95% CI 89 — 99) in the detection of arteriographically confirmed diameter stenoses of > 50%.5 However, the complexity of
assessments means disease determination is highly operator dependent, and thus requires extensive training and a detailed understanding of vascular anatomy, and there is no agreed method for reporting and documenting the findings of such scans. Therefore, a standardised method of reporting and documenting the results of DUS assessment for PAD has been developed that may be of use in clinical trials and/or in standardisation across different units/operators, to be used in addition to initial screening tools or clinical suspicion. A comparison with ABPI measurements taken as the reference standard for diagnostic performance overall and for PAD disease severity was performed.

**METHODS**

As part of an evaluation for a new diagnostic test in PAD (clinical trial registry number ISRCTN13301188), 250 participants were enrolled in primary care and divided into two groups, matched according to sex and age, based on previous diagnosis of non-critical PAD (n = 125) or the absence of PAD (n = 125) according to the General Practice (GP) register. The recruitment of these subjects took place over a 16 month period (May 2015 — September 2016) from 15 different general practices across the North East of England. As part of the main study all participants underwent baseline ankle brachial pressure index (ABPI) measurement and DUS assessment of the aorto-iliac, and infrainguinal arteries in GP practices.

Prior to commencement, the study protocol was reviewed and approved by the Newcastle & North Tyneside 1 Research Ethics Committee (14/NE/1238) and registered with the ISRCTN (ISRCTN13301188).

**Data acquisition**

Throughout the measurement process, neither practice nurses nor vascular scientists were informed of the referral diagnosis, or the findings of their counterpart. However, formal blinding was not possible.

**Ankle brachial pressure index**

ABPI was recorded by suitably trained GP practice nurses using standard techniques, using a Doppler ultrasound probe (8 MHz; Dopplex [Huntleigh Healthcare, UK]), an appropriately sized limb cuff, and a handheld mechanical sphygmomanometer. Prior to the measurement, the subject lay supine for at least 10 minutes, to allow for rest and acclimatisation.

**Duplex ultrasound**

Following ABPI measurement, the patient was scanned in a supine position with a portable US device (Sonosite M-Turbo portable ultrasound machine using imaging, colour and pulsed Doppler modes), with a 5 — 10 MHz linear array by experienced vascular scientists.

The assessment consisted of scanning along aorto-iliac, femoral, and popliteal segments, and in addition, noting the waveform shape at inguinal level, distal femoral, distal popliteal, and distal posterior and anterior tibial arteries, subsequently documenting observations on a tick box score sheet (Supplementary Appendix 1). In each of the scanned segments the following were noted: no abnormality detected; raised velocities consistent with > 50% stenosis; or a generalised narrowing/occlusion. A > 50% stenosis was indicated by at least a doubling of velocity at the stenosis (assessed by eye, rather than by direct measurement). Furthermore, for each of the distal observations the following were also noted: clear lumen with triphasic waveform; biphasic waveform with fast systolic rise time; damped waveform with long rise time; or occluded/not found.

A diagnosis of PAD was made if any of the following were identified: (1) at least one stenosis of > 50%; (2) an occlusion; or (3) general narrowing of such a degree that flow is impeded and the waveform is damped in the popliteal or distal segment. Furthermore, using the information gathered on the score sheet, a conclusion regarding anatomical/haemodynamic severity, categorised as either mild, moderate, or severe, was made on the basis of the vascular scientists’ subjective interpretation of the findings. A freeform field was also made available for additional comment, such as “significant calcification seen”.

**Data analysis**

The right and left leg per participant were handled as independent readings, comparing the DUS score directly with the corresponding ABPI value. PAD status was determined by an ABPI < 0.9, and further divided into mild, moderate, or severe using ABPI ranges illustrated in Table 1. Data were excluded if no corresponding ABPI value was obtained per DUS determination, or if the ABPI reading was > 1.4, owing to the risk of false negative results observed in calcified vessels.

All statistical analyses and data visualisation were performed using Minitab 18.1 and SciStat.com online statistical software (MedCalc Software).

**Study subjects**

Descriptive statistics were obtained, stratifying individuals into two groups based on absence, or presence of PAD in at least one leg. A Kolmogorov—Smirnov test for normality was conducted per continuous variable, and tests for statistical significance between the two groups via two tailed Student’s t test, or chi square test were performed for

<table>
<thead>
<tr>
<th>Table 1: Ankle brachial pressure index (ABPI) derived classification of severity of peripheral arterial disease of lower limbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABPI</td>
</tr>
<tr>
<td>&gt; 1.4</td>
</tr>
<tr>
<td>0.9 — 1.4</td>
</tr>
<tr>
<td>0.8 — 0.9</td>
</tr>
<tr>
<td>0.5 — 0.8</td>
</tr>
<tr>
<td>&lt; 0.5</td>
</tr>
</tbody>
</table>

Adapted from McDermott et al.
Interpretation of Cohen’s kappa to determine the inter-rater reliability

<table>
<thead>
<tr>
<th>Kappa range</th>
<th>Agreement level</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0.00</td>
<td>Poor</td>
</tr>
<tr>
<td>0.00–0.20</td>
<td>Slight</td>
</tr>
<tr>
<td>0.21–0.40</td>
<td>Fair</td>
</tr>
<tr>
<td>0.41–0.60</td>
<td>Moderate</td>
</tr>
<tr>
<td>0.61–0.80</td>
<td>Substantial</td>
</tr>
<tr>
<td>0.81–1.00</td>
<td>Almost perfect</td>
</tr>
</tbody>
</table>

Continuous and categorical variables, respectively. Note, in the above analysis, \( p \) value < 0.05 was considered to be statistically significant.

**Diagnostic performance of simplified duplex ultrasound method vs. ankle brachial pressure index**

Cross tabulation of PAD status by ABPI and DUS (according to aforementioned criteria) was performed. Subsequently, associated test sensitivity, specificity, and diagnostic accuracy (sensitivity) (prevalence) + (specificity) (1 − prevalence), using an assumed disease prevalence of 5% in the general population, were determined alongside associated 95% CIs overall, as well as per severity classification. Note, assumed disease prevalence is based on study subjects aged > 40 years in the 1999 National Health and Nutrition Examination Survey.

**Severity grading by simplified duplex ultrasound method**

Using criteria for mild, moderate, and severe disease derived by the ABPI ranges outlined in Table 1, linear weighted Cohen’s kappa was calculated to determine the inter-rater reliability with DUS derived PAD severity, alongside associated 95% CIs. Interpretation of Cohen’s kappa value throughout this study was as provided in Table 2.

**RESULTS**

**Study subjects**

From the initial participants, eight ABPI readings could not be obtained (two bilateral and six unilateral), and a further 11 measurements returned ABPI readings > 1.4 (four bilateral and three unilateral). These cases were therefore excluded from the study. In total, 244 participants were included in the final analysis, comprising of 481 legs in total, classified as either PAD present \( (n = 137) \) or absent \( (n = 344) \), determined by ABPI < 0.9 (Fig. 1).

Table 3 details the demographic data of the study subjects, and Figure 2 shows the spread of the lowest reported ABPI value of either leg per participant, stratified based on the absence of PAD in both legs, or the presence of PAD in at least one leg, as established by an ABPI < 0.9.

**Diagnostic performance of simplified duplex ultrasound method vs. ankle brachial pressure index**

Table 4 provides cross tabulation of ABPI vs. DUS determined PAD status, and Table 5 details the associated sensitivity, specificity, and diagnostic accuracy of DUS in the overall detection of ABPI determined disease, as well as specifically in mild, moderate, and severe PAD (according to ABPI ranges defined in Table 1).

Briefly, sensitivity and specificity in the detection of disease overall was 81.0% (95% CI 73.4 – 87.2) and 86.3% (95% CI 82.3 – 89.8), respectively. From mild to severe PAD, sensitivity increased from 68.8% (95% CI 53.8 – 81.3) to 89.3% (95% CI 71.8 – 97.7), with a sensitivity in the detection of ABPI determined moderate PAD similar to that of severe PAD, to 86.9% (95% CI 75.8 – 94.2). Furthermore, diagnostic accuracies ranged from 85.5% (95% CI 81.6 – 88.8) to 86.5% (95% CI 82.6 – 89.8) in ABPI determined mild, and severe disease, respectively, obtaining an overall diagnostic accuracy of 86.1% (95% CI 82.7 – 89.0).

**Severity grading by the simplified duplex ultrasound method**

Cross tabulation of ABPI determined PAD severity (in line with ABPI ranges detailed in Table 2) vs. DUS determined PAD severity obtained using the simplified DUS scoring system can be seen in Table 4. A linear weighted Cohen’s kappa value of 0.63 (95% CI 0.57 – 0.69) was obtained, indicating at least substantial agreement between the two diagnostic methods (see Table 2).

![Figure 1](image-url)
This study used a simplified and standardised report of the DUS examination, independent of any need for images or exact measurements to be recorded. Simplification and standardisation of the reporting process may prove to be time efficient, offering a useful tool for the confirmation and documentation of anatomical PAD assessment, as well as its severity following screening with first line investigations in the non-specialist/community setting where a binary result is required, or for research in PAD diagnostics, clinical trials, or when intervention is being considered.

There is an increasing call to promote standardisation of image interpretation, including, but not limited to, a reduction in the rate of diagnostic error and simpler communication of findings between colleagues. Sabel et al. evaluated the effect of structured reporting of computed tomography angiography (CTA) compared with freely dictated conventional reporting, revealing that clinicians perceive structured reporting of CTA superior in clarity, completeness, and clinical relevance. Similar findings have been reported in head and neck CTA, as well as detection of suspected pulmonary embolism via computed tomography pulmonary angiography. One of the aims of this study was to demonstrate the feasibility of such a standardised reporting system in the context of peripheral arterial DUS.

The waveform pattern recognition used in this study is in line with the waveform changes to be expected in the presence of progressing PAD (see Supplementary Appendix 2). As previously highlighted, studies that have used the above mentioned waveform changes via DUS, particularly a rise in peak systolic velocity, have reported median diagnostic sensitivities of 88% (95% CI 80–98) and specificities of 96% (95% CI 89–99) in the detection of arteriographically confirmed stenoses > 50% in diameter. Although not as significant, overall diagnostic test evaluation statistics observed in this study were near that of the comparison with the gold standard. Furthermore, a study conducted by Allen et al. previously compared colour duplex ultrasound against ABPI, observing substantial agreement between the two diagnostic tests (kappa 0.66), further supporting the present results (kappa 0.63). Interestingly in the current study, moderate agreement was established between the ABPI and DUS determined severity classification, not only verifying the diagnostic capacity, but also grading of disease by the simplified DUS
scoring system albeit modelled on the basis of the vascular scientists’ subjective interpretation of their findings. The present approach to grading anatomical PAD severity is similar to that described by Morris et al.,17 whereby an angiographic scoring system, named ANGIO score, was validated. In summary, 10 major arteries were scored 0 – 2, depending on degree of stenosis/occlusion, in which the total score correlated closely with the severity of lower limb ischaemia as measured by ABPI (Spearman’s rho = –0.33). Similarly, the Global Limb Anatomic Staging System (GLASS) defines angiographic complexity and severity of infrainguinal disease along a target arterial path.18 The authors are unaware of any similar score for DUS assessment and reporting.

The National Institute of Health and Care Excellence (NICE) guidelines state that an ABPI ratio of < 0.9 indicates the occurrence of PAD and recommends its use for diagnosis at a primary care level.19 Moreover, increasing stenosis leads to a decrease in ABPI,3,20 and can therefore be used as a means of grading disease. ABPI is limited in that it cannot localise pathology and may be unreliable in the event of vascular calcification, where falsely elevated ABPI readings (> 1.4) have been reported frequently in patients with medial calcinosis of the tibial arteries,5,7,8,21 reducing overall sensitivity in these subgroups.8 As such, the authors feel that excluding the small number of such readings in the present study was justifiable (n = 11); however, it is acknowledged that the diagnostic test accuracy statistics, including Kappa, would have been negatively affected had such readings been retained in the analysis. It is worth highlighting that while the risk of death is higher in those with an ABPI > 1.4 vs. normal ABPI ranges, the hazard ratio for death is greater in lower ABPI readings, and increases with decreasing ABPI value.22 Further, ABPI readings > 1.4 are commonly observed in patients with comorbid diabetes, chronic kidney disease, or rheumatological conditions, which invariably increase mortality risk, irrespective of the presence of PAD.5,23 Nevertheless, spectral measurement in DUS allows for determination of stenosis through waveform phascicity and derivation of peak systolic velocity, regardless of calcification,24 and is hence superior to ABPI in this regard. Admittedly, this subgroup of individuals at risk of false negative ABPI measurements would benefit from DUS determination of disease, and future confirmatory work to investigate this using the proposed standardised DUS assessment could be beneficial.

The authors are aware that there are several studies highlighting the altered lower limb haemodynamics during rest and exercise, often leading to a mismatch between resting and post-exercise ABPI readings. A study by Hammad et al. demonstrated that patients with normal resting/abnormal post-exercise ABPI readings were at higher risk of lower extremity revascularisation than those with normal resting/normal post-exercise ABPI readings.25 Therefore, one might conclude that by excluding a dynamic assessment such as this, the reference standard is prone to false negatives and thereby impacts on the validity of the results. Unfortunately, however, post-exercise ABPI was not measured in the wider study (NOTEPAD), which was guided by current NICE guidelines. Furthermore, it is worth mentioning that while ABPI can reliably detect significant stenosis,5,8,26 the diagnostic power of ABPI in mild disease is debatable.8,26 Given that this study was carried out in a primary care based setting, more complex imaging, which would act as a true independent reference “gold” standard, could better detect “mild stenosis”. Future work could include a comparison of DUS determined disease (using the standardised simplified DUS examination) with a more robust reference standard, such as CTA or magnetic resonance angiography. Additionally, following validation with angiographic methods, it is proposed that the standardised DUS system could be used to follow up patients with remaining clinical suspicion of PAD, despite normal ABPI ranges.

Finally, in the light of the various limitations posed by ABPI, there are several novel tools, including toe flexion near infrared spectroscopy (NIRS)27 and multisite photoplethysmography,28 evolving in the field of vascular optics offering potential alternatives for PAD screening in primary care/community settings, the former presenting extremely promising diagnostic test accuracy when compared with echo colour Doppler, even in the presence of diabetes. The proposed standardised DUS scoring system could be used as an adjunctive tool for confirmation of diagnosis and subsequent severity classification. Further, the diagnostic performance of the standardised DUS assessment could be compared with novel screening tools to further validate these findings.

**Conclusion**

Duplex ultrasound is a non-invasive method of arterial mapping, proven to diagnose reliably PAD, grade severity accordingly, and localise stenosis as a means of intervention planning. The standardised DUS method used in this study performed exceptionally well with regard to overall diagnosis of ABPI determined disease (sensitivity 81.0% [95% CI 73.4 – 87.2%]; specificity 86.3% [95% CI 82.3 – 89.8%];

**Table 5. Diagnostic test accuracy statistics of Duplex ultrasound vs. ankle brachial pressure index determined status of peripheral arterial disease (PAD) of lower limbs in 481 limbs of 244 patients**

<table>
<thead>
<tr>
<th>PAD status</th>
<th>Sensitivity (95% CI) – %</th>
<th>Specificity (95% CI) – %</th>
<th>Accuracy (95% CI) – %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>81.0 (73.4–87.2)</td>
<td>86.3 (82.3–89.8)</td>
<td>86.1 (82.7–89.0)</td>
</tr>
<tr>
<td>Mild</td>
<td>68.8 (53.8–81.3)</td>
<td>85.5 (81.6–88.8)</td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>86.9 (75.8–94.2)</td>
<td>86.4 (82.6–89.6)</td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>89.3 (71.8–97.7)</td>
<td>86.5 (82.6–89.8)</td>
<td></td>
</tr>
</tbody>
</table>

*CI = confidence interval.*
accuracy 86.1% [95% CI 82.7 — 89.0]), in addition to substantial agreement with ABPI determined severity of disease (kappa 0.63, 95% CI 0.57 — 0.69), offering a potential tool for confirmation of diagnosis and anatomic/haemodynamic severity assessment, when used in conjunction with ABPI and/or evolving non-invasive screening tools. Moreover, modern ultrasound devices allow for portability without loss of image quality, making their use in clinical practice increasingly realistic. A significant limitation of this study concerns the exclusion of ABPI readings > 1.4, along with corresponding DUS findings, owing to the risk of false negative results in the event of vascular calcification. Future confirmatory work could include a comparison of DUS determined disease using the standardised scoring system, along with gold standard methods of non-invasive angiography, novel diagnostic tools such as toe flex NIRS, as well as the investigation of clinically suspected PAD in patients with normal ABPI ranges and/or those with suspected vascular calcification, in whom measurement of ABPI is unreliable.

CONFLICT OF INTEREST
Between 2014 and 2018, Professor John Allen was the Chief investigator of a National Institute for Health and Care Research i4i funded grant (Invention for Innovation, “Innovative photoplethysmography technology for rapid non-invasive assessment of peripheral arterial disease in primary care”, II-C1-0412-20003). There are no other potential conflicts of interest to report.

FUNDING
This report is independent research funded by the National Institute for Health Research (Invention for Innovation, “Innovative photoplethysmography technology for rapid non-invasive assessment of peripheral arterial disease in primary care”, II-C1-0412-20003). The views expressed in this publication are those of the author(s) and not necessarily those of the National Health Service, the National Institute for Health Research, or the Department of Health and Social Care.

ACKNOWLEDGEMENTS
We would like to thank all practices and participants for contributing to the study, as well as to those of the wider NOTEPAQ team, specifically Lesley Wilson, who contributed significantly to study design and data collection.

APPENDIX A. SUPPLEMENTARY DATA
Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejvs.2022.04.013.

REFERENCES
A 67 year old man presented with acute right lower extremity ischaemia (Rutherford stage IIb) following total hip replacement. Digital subtraction angiography demonstrated occlusion of the femoral bifurcation (red arrow; A). During open surgery, an extensive thermal injury related to the prosthesis implantation was seen on the arterial wall and in the surrounding soft tissue furthermore, a fistula (white arrow) was detected to the hip joint. The arterial bifurcation was replaced using autologous saphenous vein (B). The patient had an uneventful post-operative recovery and was discharged for rehabilitation with a patent vascular reconstruction.